Manuel d'utilisation du prototype d'étiquetage et première expérimentation (fin 2008)

Julien Lesbegueries

1 Installation

Le prototype est constitué d'une archive Java (fichier jar), d'un fichier de propriétés (OJB.properties) et d'un dossier contenant les fichiers de *mapping* (mapping). Il nécessite aussi la connexion à un serveur de base de données géographique (Postgres/PostGIS). Si la connexion se fait avec le serveur Postgres du LSIIT (FODOMUST) il faut demander l'ouverture du port 5432 et donner la plage d'adresses IP (des machines clientes) aux administrateurs réseaux du LSIIT.

1.1 Configuration

Il faut commencer par copier le jar, le fichier OJB.properties et le dossier mapping dans un dossier X (de chemin absolu /chemin/absolu/X).

La seule configuration à effectuer consiste à modifier la variable « repositoryFile »qui se trouve dans le fichier OJB.properties :

```
repositoryFile=/chemin/absolu/X/mapping/ojb1/lesbegu/repository.xml
```

1.2 Lancement

Le script lancement_prototype.sh lance le prototype (sous Unix/Linux assimilés). Il contient la commande suivante que l'on peut taper sur n'importe quelle plate-forme, depuis le dossier X :

```
java  -Xmx400m -DOJB.properties=/chemin/absolu/X/OJB.properties
      -jar PROTO_MARQUAGE_ILOTS_fat.jar
```

2 Utilisation

L'interface de départ (Fig. 1) propose le menu suivant.

– Etiqueter : Action par défaut une fois que l'on a créé un compte utilisateur.

– Ajouter un utilisateur : Première action à effectuer quand on lance le programme pour la première fois.
– Progression de l'annotation : Permet de comparer le pourcentage d'annotation effectué par chaque utilisateur.

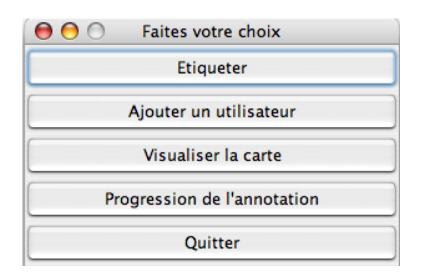


Fig. 1 – Affichage au départ.

2.1 Étiqueter

L'interface d'étiquetage (Fig. 2) propose de choisir

– Le mode d'annotation : normal ou confiance. Le mode « normal » oblige l'utilisateur à choisir catégoriquement une classe d'annotation par îlot. Le mode « confiance » laisse à l'utilisateur la possibilité de graduer son choix (de 1 à 4[1]) et de choisir plusieurs classes pour un îlot.
– Le choix de la ville,
– Le choix de la zone dans cette ville,
– Le mode d'affichage de la carte (îlot et ses voisins ou zone complète).

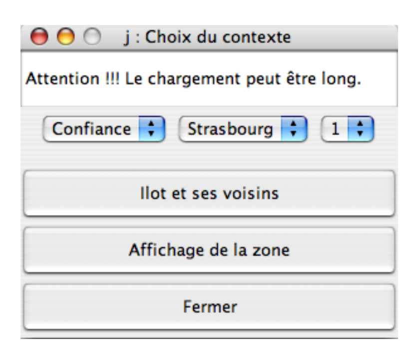


Fig. 2 – Interface d'étiquetage.

[1]par exemple : non, probablement que non, probablement que oui, oui ?

2.1.1 Îlot et ses voisins

Dans ce mode d'affichage, l'îlot à annoter se trouve en jaune et il est centré sur l'écran. Il est visualisé avec les îlots ainsi que les objets micro voisins (Fig. 3).

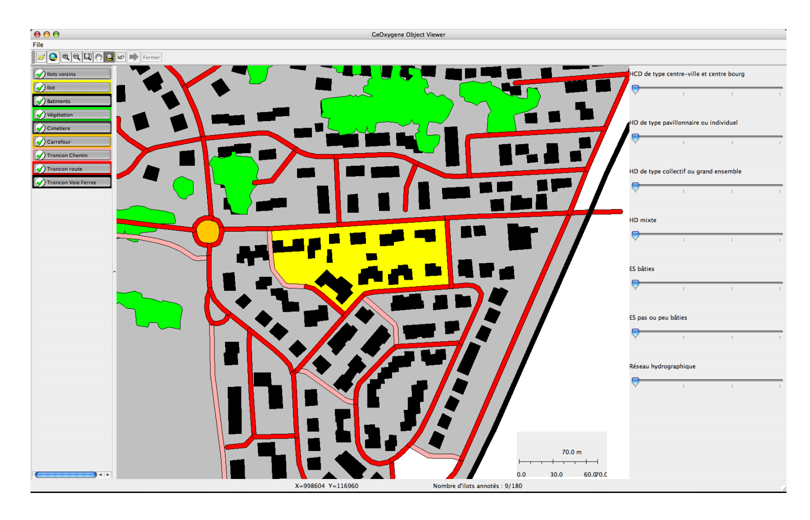


Fig. 3 – Visualisation de l'îlot à annoter et des objets micro voisins (bâtiments, tronçons de route, etc.).

Certains boutons sont disponibles dans le menu et en particulier :

– est le bouton sélectionné par défaut. Il permet de zoomer sur une partie de la carte (en dessinant un rectangle à la souris).

– permet de visualiser l'ensemble des couches présentes sur la carte.

– permet de se déplacer dans la carte.

– Id? permet d'obtenir les informations présentes dans la base de données pour les objets cliqués.

Le panneau latéral gauche permet de voir quelles sont les couches affichées. En cas de problème d'occlusion, un clic droit ouvre un menu contextuel qui permet de rendre invisible ou de changer l'ordre des couches.

2.1.2 Affichage de la zone

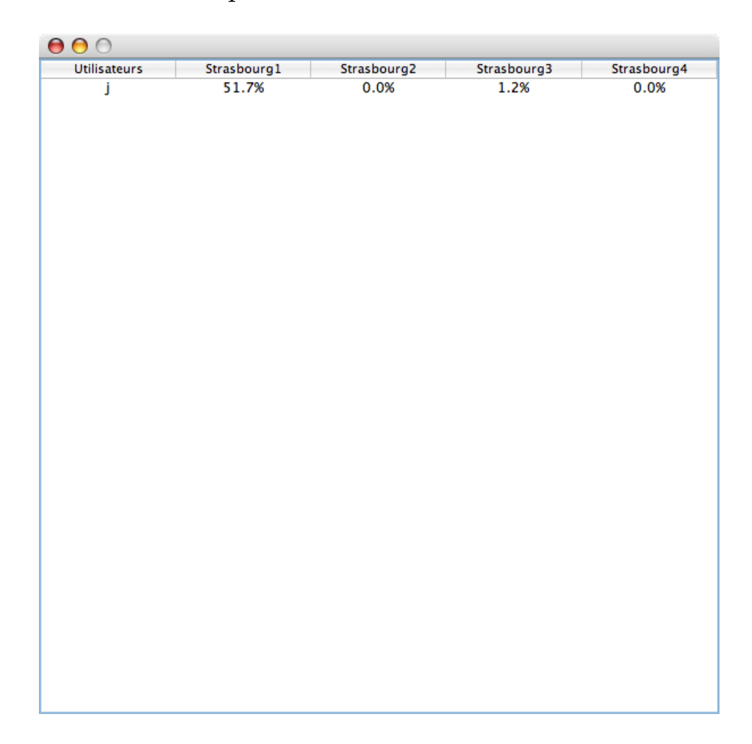
Dans ce mode, la zone complète est affichée avec 2 couches : la première pour les îlots annotés, la deuxième pour ceux qui ne le sont pas encore. Il est alors possible de procéder à l'annotation (avec un clic droit sur l'îlot).

3 Procédure d'annotation

Dans cette section est décrite la procédure d'annotation, décidée pour le moment (septembre 2008). Les règles de départ sont :

- Un compte utilisateur par personne,
- Annotation dans le contexte « confiance »,
- Annotation dans le mode de visualisation « Îlot et ses voisins ».

L'annotation se fait sur toutes les zones disponibles. Il est demandé de limiter l'utilisation du bouton Id? et d'utiliser autant que possible uniquement l'information présente à l'écran pour annoter. Le bouton « Progression de l'annotation » permet de voir le pourcentage de progression pour chaque zone, pour chaque utilisateur (Fig. 4). Il serait intéressant d'avoir un maximum de recoupements pour les annotations d'îlot. Ce tableau permet donc d'aiguiller les utilisateurs sur les prochaines zones à annoter.



Utilisateurs	Strasbourg1	Strasbourg2	Strasbourg3	Strasbourg4
j	51.7%	0.0%	1.2%	0.0%

Fig. 4 – Tableau d'affichage de la progression d'annotation des îlots.

La procédure d'annotation est la suivante :

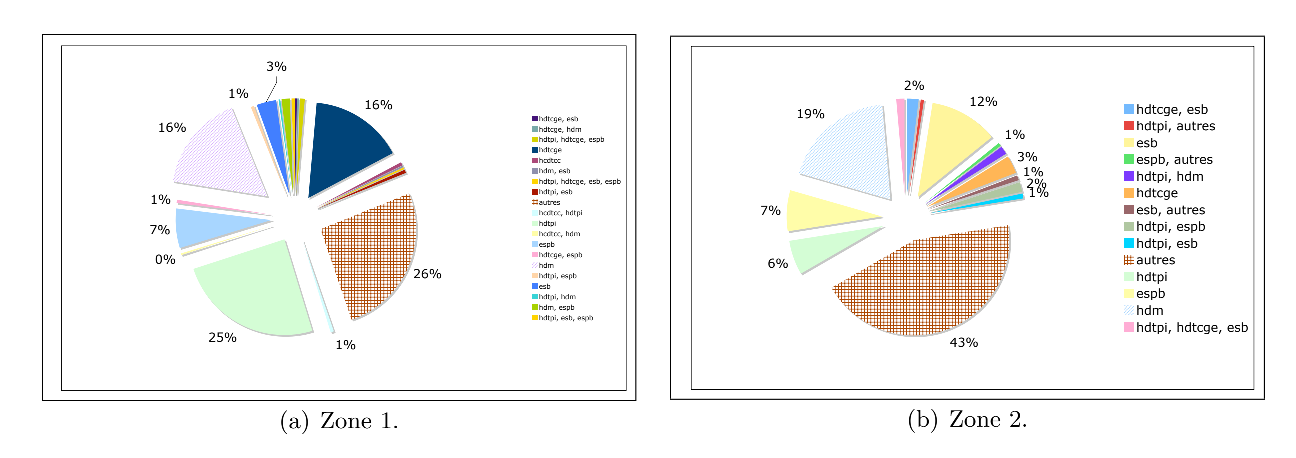
1. Créer un compte utilisateur,
2. S'identifier (au niveau de l'étiquetage),
3. Visualiser l'îlot (en jaune) et procéder à l'annotation,

4. Appuyer sur le bouton ➡ pour passer à l'îlot suivant ou sur le bouton « Fermer » pour fermer le mode d'affichage « îlot et ses voisins » . **Attention : une fois que l'on a cliqué sur ➡, il est interdit de fermer l'application sans avoir annoté.**

4 Bilan du premier étiquetage

L'étiquetage a été effectué par les experts de Strasbourg (Anne Puissant, Grzegorz Skupinski et Christiane Weber) sur les zones de Strasbourg.

4.1 Analyse des résultats



(a) Zone 1.

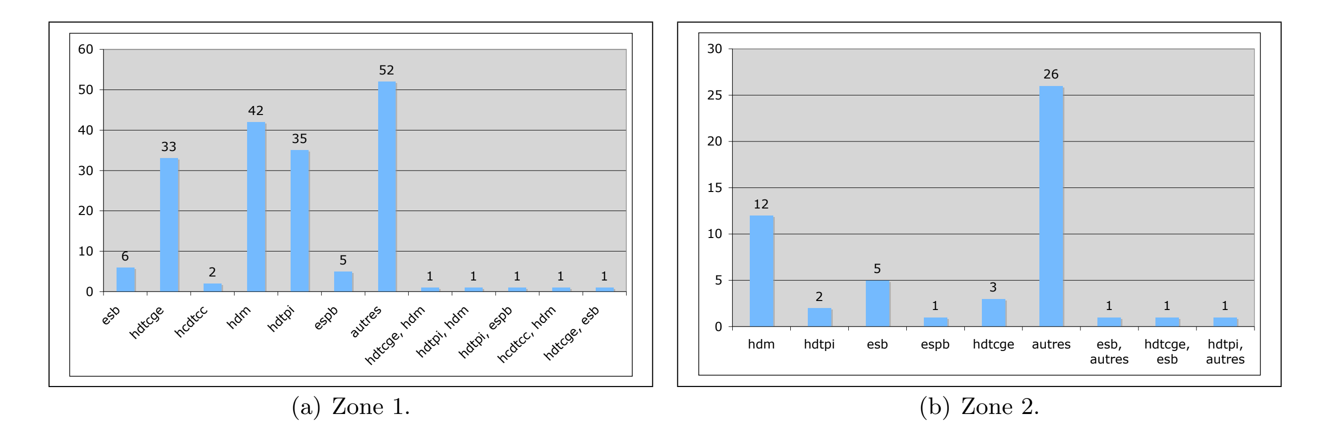
(b) Zone 2.

FIG. 5 – Répartition des classes prédominantes.

La figure 5 montre la répartition des classes prédominantes par zone, c'est-à-dire, les classes ayant remportées le plus de votes lors de l'étiquetage des 3 experts[2]. On peut voir que les classes *autres* (rouge quadrillé) et *hdm* (bleu rayé) sont quasiment majoritaires.

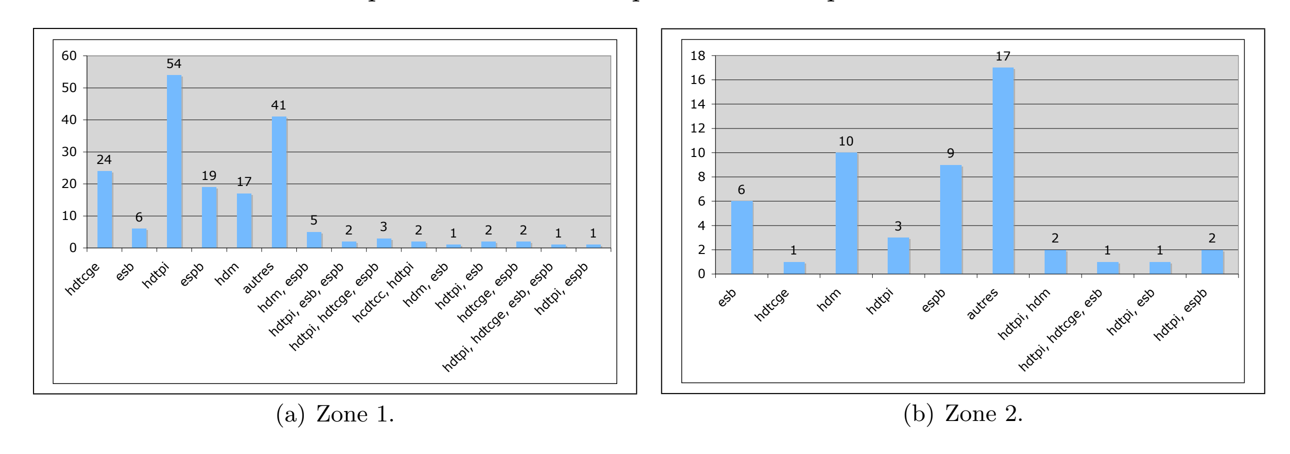
Les figures 6, 7, 8 montrent respectivement les résultats détaillés d'Anne, de Grégoire et de Christiane. Ces résultats montrent la quantité importante d'îlots classés *autres* et *hdm* posant problème (Le tableau 1 détaille aussi les pourcentages d'îlots étiquetés dans ces 2 classes). Ils montrent aussi qu'une certaine quantité d'îlots ont été étiquetés par plusieurs classes à la fois.

[2]Étant donné que la classification unique était difficile pour les experts, les diagrammes de la figure 5 présentent la liste des combinaisons de classes utilisées.



(a) Zone 1. (b) Zone 2.

Fig. 6 – Répartition des classes prédominantes pour Anne.



(a) Zone 1. (b) Zone 2.

Fig. 7 – Répartition des classes prédominantes pour Grégoire.

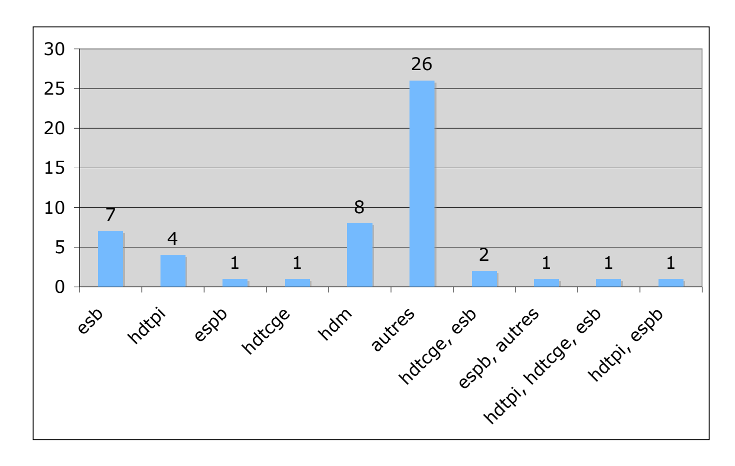


Fig. 8 – Répartition des classes prédominantes pour Christiane.

Expert	Zone 1		Zone2		Total	
	autres	hdm	autres	hdm	autres	hdm
Anne	28,88%	23,33%	50%	23,08%	39,44%	23,21%
Grégoire	22,77%	9,44%	32,69%	19,23%	27,73%	14,34%
Christiane	-	-	50%	15,38%		
Moyenne	25,83%	16,39%	44,23%	19,23%		

TAB. 1 – Pourcentage d'étiquetage en « autres » et« hdm ».

4.2 Problèmes rencontrés

Bord de ville Les experts ont eu des difficultés à étiqueter les îlots de bord de ville (figure 10). Une proposition est donc faite de les enlever de la base d'étiquetage (et d'apprentissage).

Habitat Discontinu Mixte La classe *hdm* a introduit une confusion dans l'étiquetage et des divergences apparentes entre les experts. Un habitat discontinu mixte étant composé d'habitat discontinu de type pavillonnaire et individuel ainsi que d'habitat discontinu de type collectif et grands ensembles, les experts ont étiqueté un même îlot comme étant *hdm* ou *hdtpi, hdtcge*. La figure 9(a) montre un exemple d'îlot pour lequel l'étiquetage actuel pose problème et, en particulier, l'existence de la classe *hdm*. En effet, pour cet îlot :

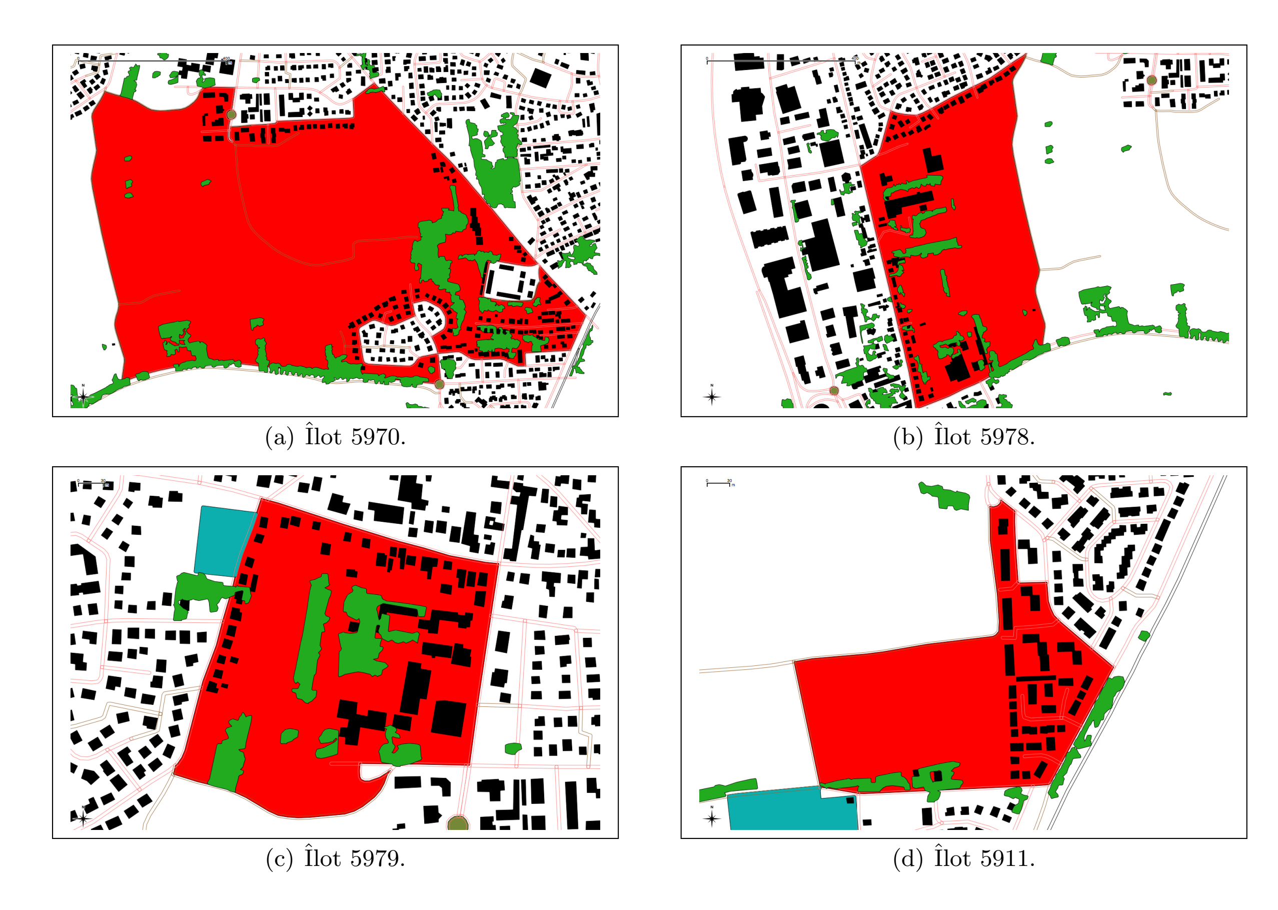
- Anne a choisi « Habitat Discontinu Mixte » avec une gradation de 4 sur 4, indiquant une confiance forte.
- Grégoire a choisi 3 classes (4/4 pour les 3) : « Habitat discontinu de type pavillonnaire ou individuel » , « Habitat discontinu de type collectif ou grand ensemble » et « Emprises spécialisées pas ou peu bâties », car il considère que ses 3 caractéristiques sont présentes dans cet îlot.

  D'autres îlots ont subi le même genre d'étiquetage (figures9(b), 9(c), 9(d)).

  Le problème est donc que les curseurs ne sont pas utilisés de la même manière par les différents experts.

Carrefours et espaces nus Certains carrefours et espaces nus ont été classés *autres* alors que le RT 2.1 indique qu'il s'agit d'emprises spécialisées bâties (ou pas ou peu bâties).

Les figures 11,12 montrent quelques exemples d'îlot ayant posé des problèmes pour l'étiquetage.



(a) Îlot 5970.

(b) Îlot 5978.

(c) Îlot 5979.

(d) Îlot 5911.

Fig. 9 – Exemple d'îlot étiqueté de 2 manière différentes par les experts.

5 Nouvelle procédure

La nouvelle procédure tient compte du fait que les îlots ne sont pas découpés comme le voudraient les experts. Ceux-ci ne veulent pas assigner une classe unique par îlot. Les curseurs correspondent donc maintenant à un degré de certitude par rapport à la présence ou non d'une classe dans un îlot. Par exemple, l'expert peut étiqueter un îlot comme étant *hdtcge* et *hdtpi*, ce qui signifie que l'îlot contient des habitations de type collectives (grands ensembles) ET des habitations de type pavillonnaire (individuelles).

De fait, la classe *hdm* (mixte) peut disparaître, ainsi que la classe *autres* qui était utilisée pour indiquer que l'îlot posait problème[3]. À la demande des experts, une nouvelle classe « réseaux de communication » *(rc)* a été ajoutée.

[3]Des commentaires peuvent maintenant être laissés pour chaque îlot à étiqueter, pour chaque expert, en cas de problème.

Résumé de la procédure

- Étiquetage multi-classes (dans un îlot, il peut y avoir plusieurs classes),
- La gradation des curseurs indique le degré de certitude de l'étiquetage (de 1, pas du tout, à 4, sûr).
- Les étiquetages réalisés durant la première phase sont sauvegardés et proposés à l'expert (les curseurs *hdm* et *autres* étant grisés) afin d'accélérer le processus d'étiquetage.

Annexe



Fig. 10 – Îlots marqués comme « autres » (en rouge) dans la zone 1 et correspondant aux bords de ville.

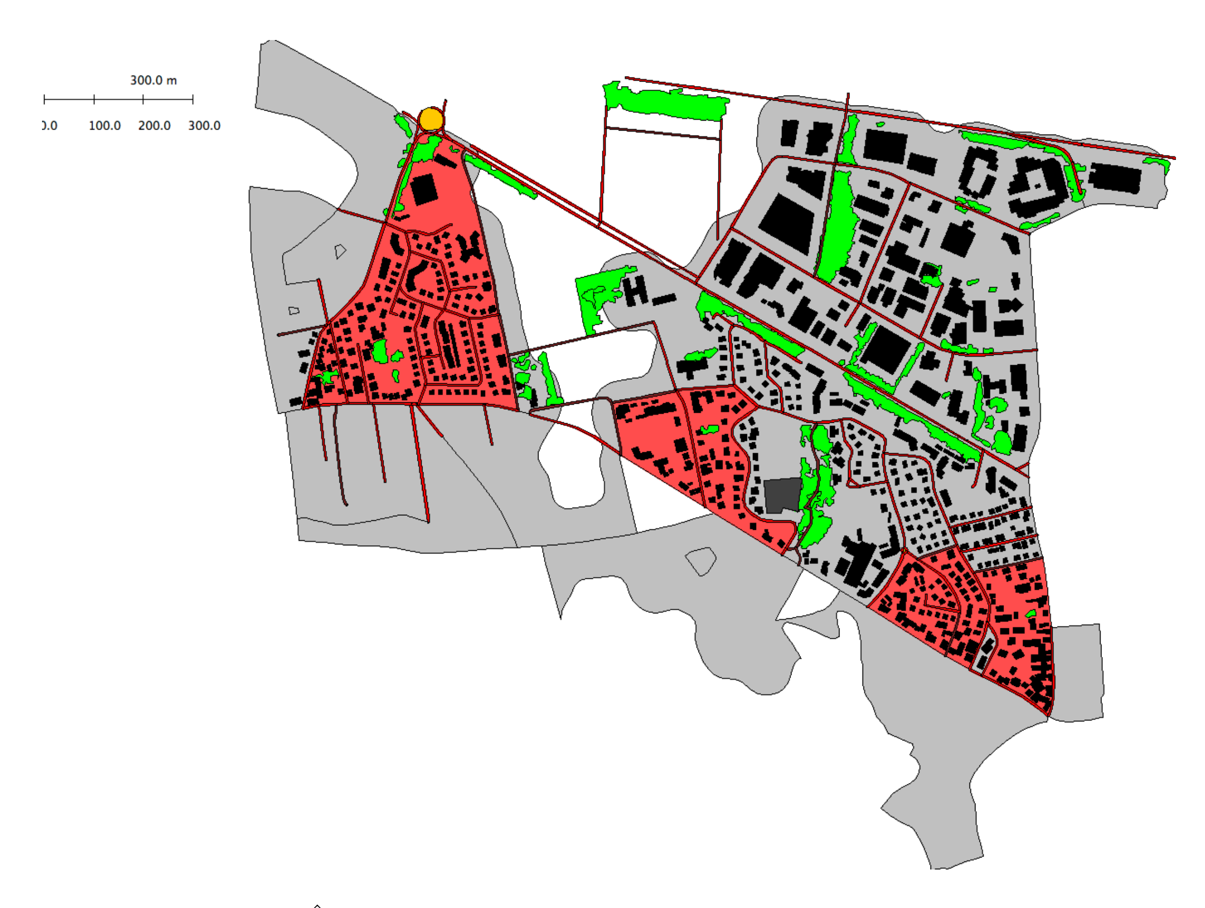


Fig. 11 – Îlots marqués comme « hdm » (en rouge) dans la zone 2.



Fig. 12 – Îlots marqués comme « autres » (en rouge) dans la zone 2.